An Exascale Perspective 2010-2015

Justin Rattner
Intel Senior Fellow and CTO (retired)

November 16, 2015

Outline

- Exascale view in 2010
- Exascale view in 2015
- Moore’s Law Obituary
- NTV and Dynamic Power Control
- The Exascale Elephant

The Exceptional Challenges of Exascale Computing

Justin Rattner
Chief Technology Officer
Intel Corporation

January 19, 2010

Gettin' to Exascale Ain't Gonna be Easy

Justin Rattner
Chief Technology Officer
Intel Corporation

The Path Forward

Research Needed to Achieve Exascale Performance

- Extreme voltage scaling to reduce core power
- More parallelism 10x – 100x to achieve speed
- Re-architecting DRAM to reduce memory power
- New interconnect lower power and distance
- NVM to reduce disk power and accesses
- Resilient design to manage unreliable transistors
- New programming tools for extreme parallelism
- Applications built for extreme parallelism

HPC Node Architecture

What it is, and what it should be…

Shekhar Borkar
Intel Corporation
July 14, 2015

This research was, in part, funded by the U.S. Government, DOE and DARPA. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

Top HPC Challenges

(from Shekhar Borkar’s July, 2015 ISC Talk)

1. **System Power & Energy**
2. **New, efficient, memory subsystem**
3. **Extreme parallelism**
   - Data locality,
   - Programmability
4. **New execution model**
   - Self awareness
   - Introspection
5. **Resiliency for system reliability**
6. **System efficiency & cost**

NATIONAL STRATEGIC COMPUTING INITIATIVE
July 29, 2015

EXECUTIVE ORDER

- - - - - - -

CREATING A NATIONAL STRATEGIC COMPUTING INITIATIVE

By the authority vested in me as President by the Constitution and the laws of the United States of America, and to maximize benefits of high-performance computing (HPC) research, development, and deployment, it is hereby ordered as follows:

The NSCI is a whole-of-government effort designed to create a cohesive, multi-agency strategic vision and Federal investment strategy, executed in collaboration with industry and academia, to maximize the benefits of HPC for the United States.

https://www.whitehouse.gov/the-press-office/2015/07/29/executive-order-creating-national-strategic-computing-initiative
https://www.whitehouse.gov/sites/default/files/microsites/ostp/nsci_fact_sheet.pdf

Just How Dead is Moore’s Law?

Moore's Law has Died More than Once

Moore’s Law for Si-Gate CMOS Ended at 65 nm

(A 40-Year Journey)

Moore’s Law for Planar Transistors Ended at 32nm

Moore's Law Began for 3D Transistors at 22nm

What's Next for Moore's Law?
For Intel, III+V = 10nm QWFETs

April 21, 2015 by David Kanter (realworldtech.com)

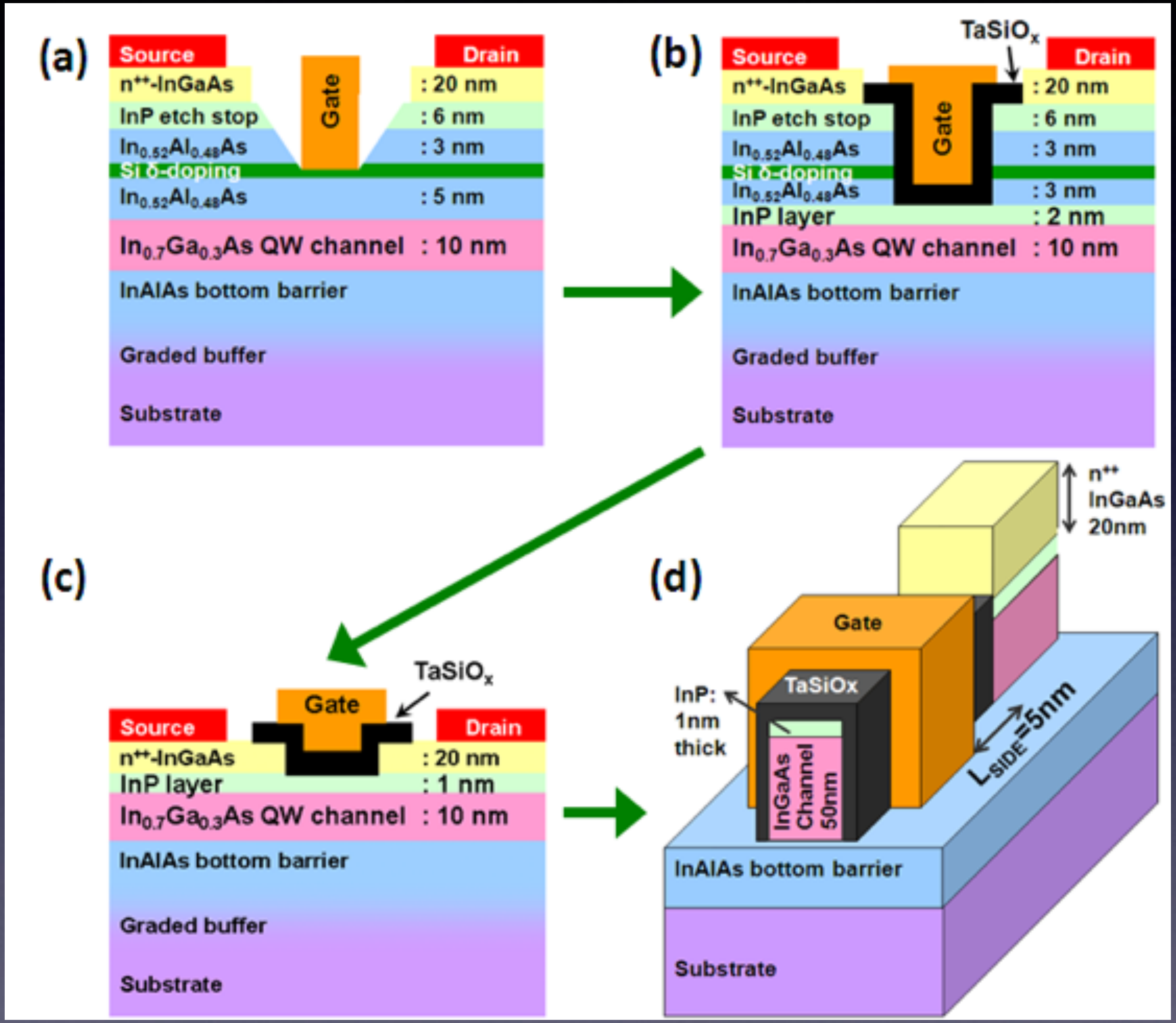
"The industry will adopt Quantum Well FETs (QWFETs) that use a fin geometry and high-mobility channel materials to achieve excellent transistor performance at nominal operating voltages around 0.5V (compared to roughly 0.7V for FinFETs)

The industry will adopt III-V compound semiconductors (most likely In0.53Ga0.47As, alternatively InSb) for the n-type QWFET channel

The industry will adopt strained Germanium (most likely) or III-V materials (as an alternative) for the p-type QWFET channelIntel will adopt QWFETs at the (mo15 or early 2016 (alternatively at 7nm in 2017 or 2018)

Intel will probably co-integrate conventional transistors and QWFETs, it is less likely (but possible) that the company will use separate substrates that are packaged together to optimize cost."

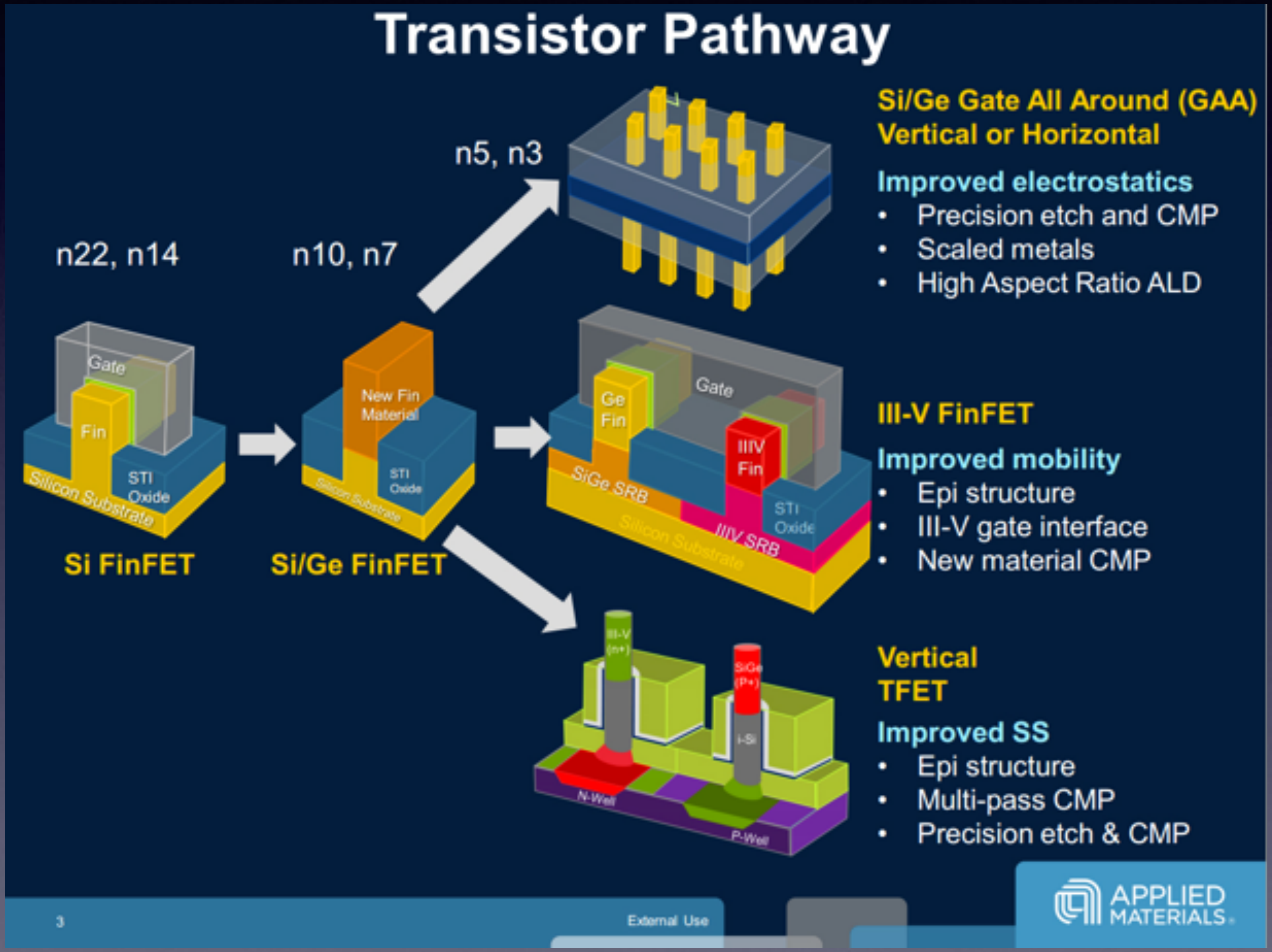
QWFET Evolution



Non-planar, multi-gate InGaAs quantum well field effect transistors with high-k gate dielectric and ultra-scaled gate-to-drain/gate-to-source separation for low power …

M Radosavljevic, G Dewey… - … (IEDM), 2010 IEEE

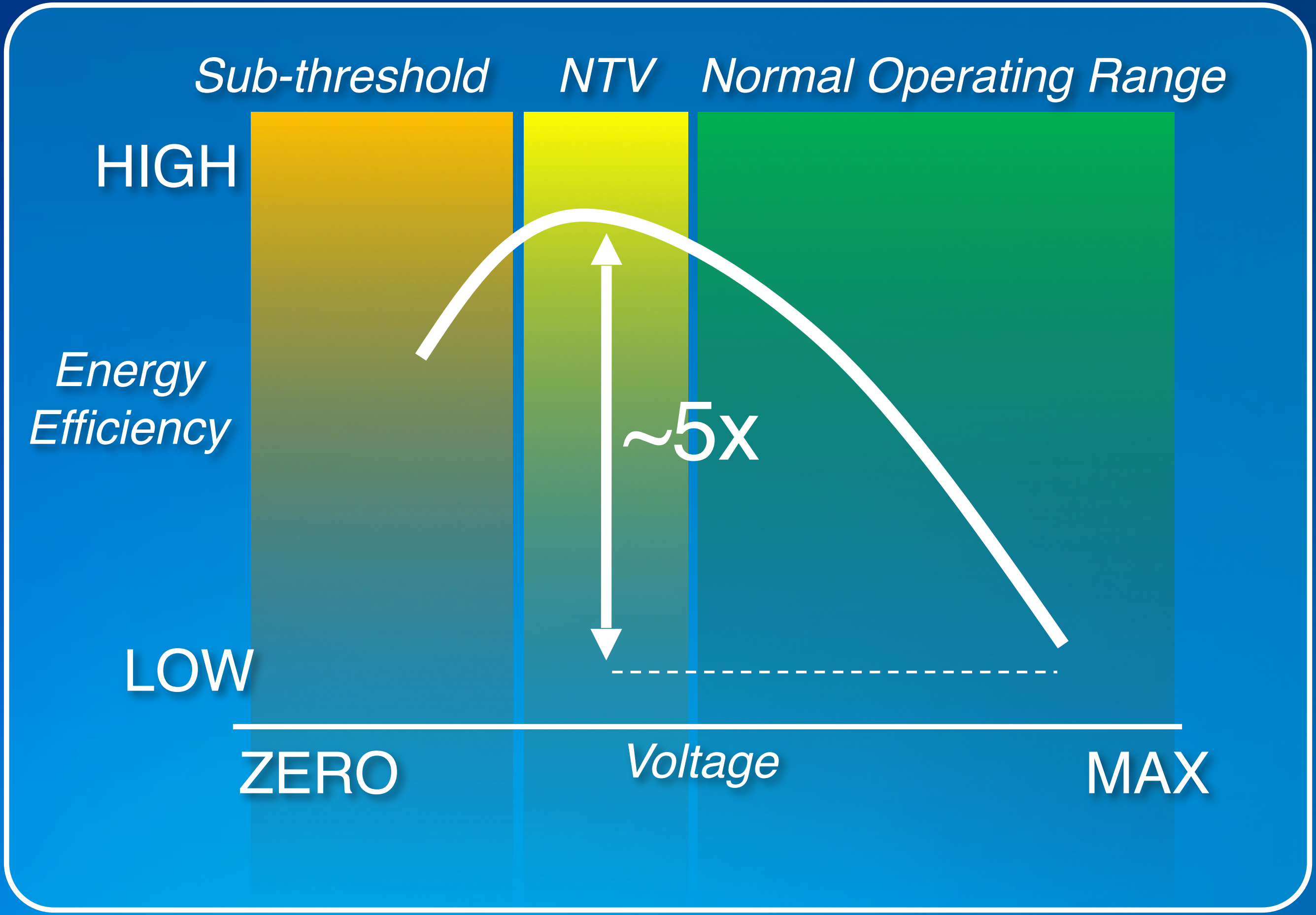
Equipment Industry Prepares for III-V FinFETs



Adam Brand, "Applied Materials Precision Materials to Meet FinFET Scaling Challenges Beyond 14nm," Semicon West, July 2013

Benefits of Near Threshold Voltage Operation

Peak energy efficiencies at NTV and fine-grain power management



Potential for…

- More always-on / instant wake devices
- Intelligent everyday devices with battery/solar powered CPUs
- Longer battery lives for mobile computing
- Scalable many-core chips for the datacenter
- Meeting extreme-scale compute challenges

The Elephant in the Room

Current partnerships with vendors
Fast and Design Forward Programs

Fast Forward Program – *node technologies*

- Jointly funded by SC & NNSA
- Phase 1: Two year contracts, started July 1, 2012, Phase 2: Two year contracts, starting Fall 2014: IBM, Cray, AMD, NVIDIA, Intel ($64M / $100M)

Project Goals & Objectives

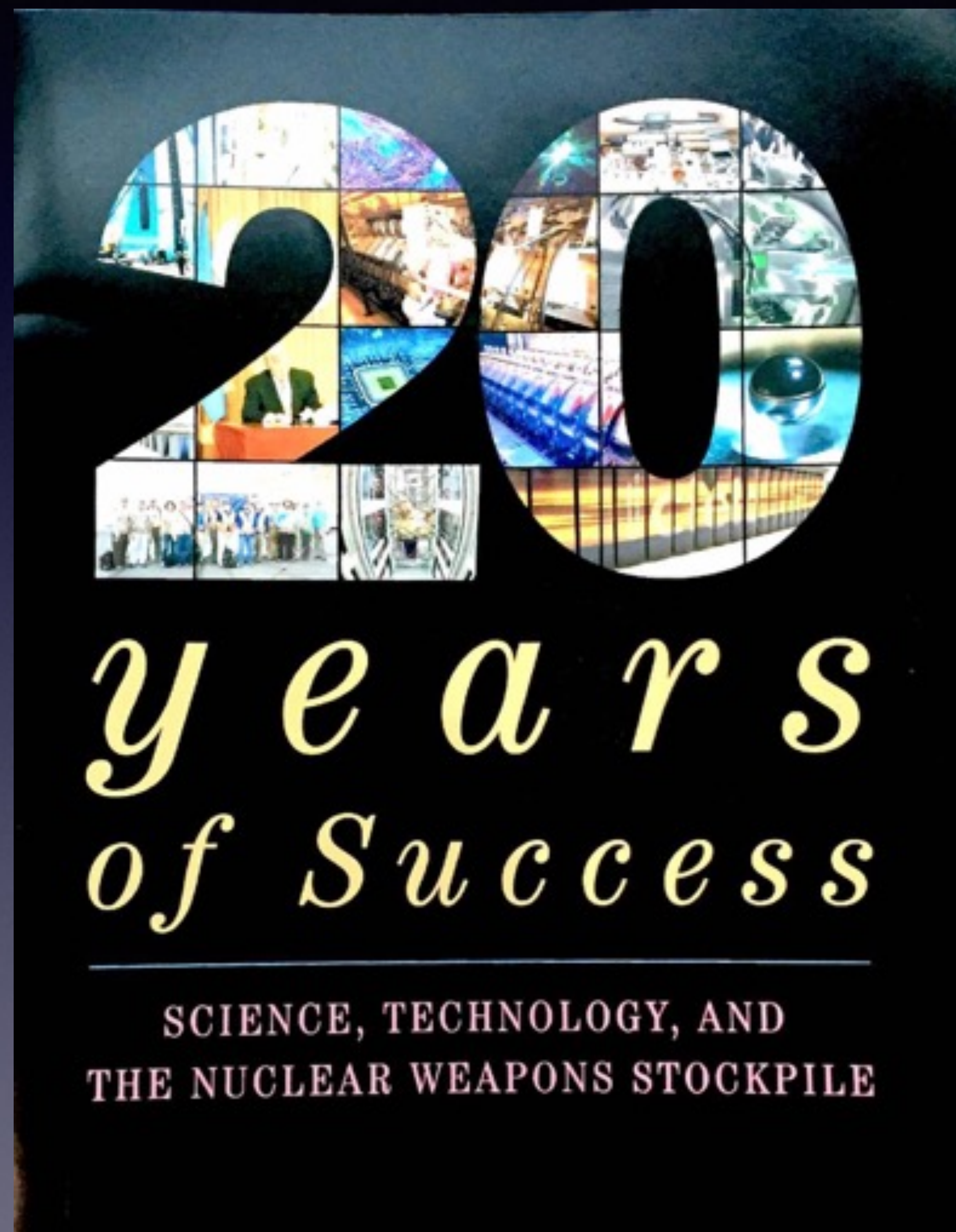
- Initiate partnerships with multiple companies to accelerate the R&D of critical node technologies and designs needed for extreme-scale computing.
- Fund technologies targeted for productization in the 5–10 year timeframe.

Design Forward Program – *system technologies*

- Jointly funded by SC & NNSA
- Phase 1: Two year contracts, started Fall 2013, Phase 2: Two year contracts. Starting Winter 2015: Cray, AMD, IBM, Intel ($23M / $10M)

Project Goals & Objectives

- Initiate partnerships with multiple companies to accelerate the R&D of interconnect architectures and conceptual designs for future extreme-scale computers.
- Fund technologies targeted for productization in the 5–10 year timeframe.

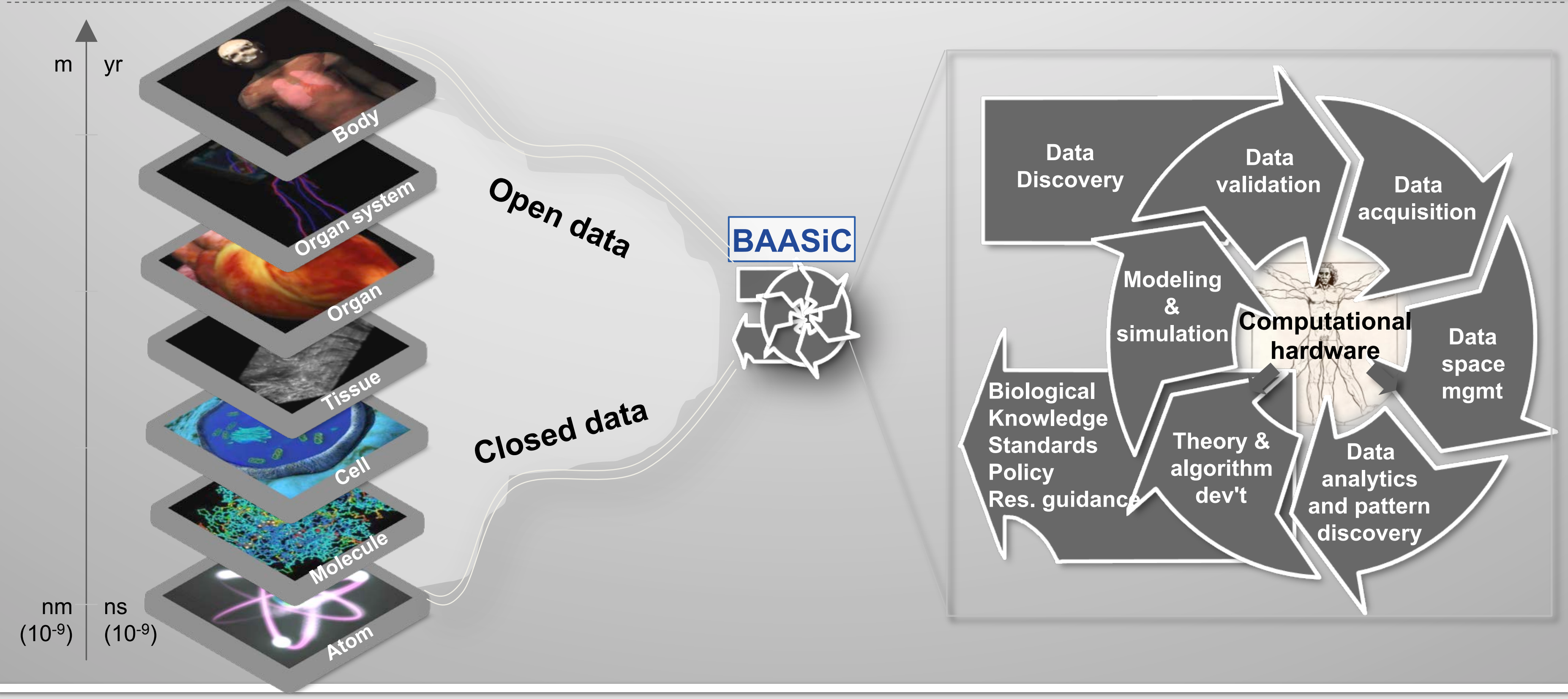

20
years
of Success
SCIENCE, TECHNOLOGY, AND
THE NUCLEAR WEAPONS STOCKPILE

Exascale Continues to Lack a Strategic Imperative

Biological Applications of Advanced Strategic Computing (BAASiC)

BAASiC will drive **disruptive improvement** in biomedical value chain through **predictive biology**

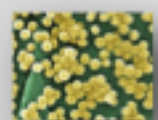
BAASiC Premise	Biology is an interconnected system; we are nearing the limits of siloed approaches
BAASiC Vision	Revolutionary computational approach to improve human health and security through
Path to vision	An integrated biological data ecosystem guiding experiment, theory, and simulation



Why now?

There is both the need and the means

There are pressing crises in biology impacting US and beyond...



Continual rise of antimicrobial resistance, Enterovirus D68, and Ebola virus
- Decreasing pharmaceutical productivity

Drug R&D at or below cost of capital
- Combined with rising cost of bringing new drugs to market

Emerging and engineered pathogens pose unprecedented asymmetric threats
- Democratization of biology
- Continued life sciences advancements

Digital revolution has missed the biomedical sector
- Advances in other sectors are computational driven

...and advances that allow us to address them

Biology

Million-fold decrease in genetic sequencing cost
- Allow association studies linking genetics to disease
- Personalized therapy based on individual genetics within reach

Massive clinical data explosion
- Need to marry disparate data sets with new computational tools

Computation

Exponential growth in high-performance computational power
- Growth globally, not just US

Rise and commercial-based Cloud compute services
- Continual innovation allow democratization of advanced computing capabilities

Exascale Five Years On

Good News: Moore's Law is Alive and Well

Bad News: Still No Strategic Imperative

Thank You!

Back-up Content

Commercial Relevance of Exascale

- Relevant and time-critical
  - Extreme voltage scaling
  - Power-reduced, bandwidth enhanced DRAM
  - Power-reduced, latency enhanced storage
  - Low-cost, low-power photonic interconnects
  - Resilient design to manage unreliable transistors
- Less Relevant or time-critical
  - Extremely parallel systems (millions/billions of threads)
  - New programming tools for extreme parallelism
  - Applications built for extreme parallelisma